COMPUTERS & SECURITY 73 (2018) 235–248

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cose

Computers & Security

Toward a more dependable hybrid analysis of android malware using aspect-oriented programming

Aisha I. Ali-Gombe [a,*], *Brendan Saltaformaggio* [b], *J. "Ram" Ramanujam* [c], *Dongyan Xu* [d], *Golden G. Richard III* [c]

[a] *Department of Computer and Information Science, Towson University, RM 447, 7800 York Road, Towson, MD 21252, USA*
[b] *School of Electrical and Computer Engineering, Georgia Institute of Technology, Klaus Advanced Computing Building, 266 Ferst Dr NW, Atlanta GA 30332, USA*
[c] *Center for Computation and Technology, Louisiana State University, 2027-C Digital Media Center, Baton Rouge, LA 70803, USA*
[d] *Department of Computer Science, Purdue University, 305 N. University Street, West Lafayette, IN 47907, USA*

ARTICLE INFO

Article history:
Received 24 March 2017
Received in revised form 2 November 2017
Accepted 5 November 2017
Available online 21 November 2017

Keywords:
Hybrid analysis
Bytecode weaving
Instrumentation
Dynamic execution
Android
Malware
Dataflow

ABSTRACT

The growing threat to user privacy by Android applications (app) has tremendously increased the need for more reliable and accessible analysis techniques. This paper presents *AspectDroid*[1]—an offline app-level hybrid analysis system designed to investigate Android applications for possible unwanted activities. It leverages static bytecode instrumentation to weave in analysis routines into an existing application to provide efficient dataflow analysis, detection of resource abuse, and analytics of suspicious behaviors, which are then monitored dynamically at runtime. Unlike operating system or framework dependent approaches, *AspectDroid* does not require porting from one version of Android to another, nor does it depend on a particular Android runtime, making it a more adaptable and easier to use technique. We evaluate the strength of our dataflow algorithm on 105 apps from the DroidBench corpus, with experimental results demonstrating that *AspectDroid* can detect tagged data with 94.68% accuracy. Furthermore, we compare and contrast the behavioral patterns in 100 malware samples from the Drebin dataset (Arp et al., 2014) and 100 apps downloaded from Google Play. Our results showed more traces of sensitive data exfiltration, abuse of resources, as well as suspicious use of programming concepts like reflection, native code, and dynamic classes in the malware set than the Google Play apps. In terms of runtime overhead, our experiments indicate *AspectDroid* can comprehensively log relevant security concerns with an approximate overhead of 1 MB memory and a 5.9% average increase in CPU usage.

1. Introduction

Android malware and over-privileged applications are well-known for privacy violations and data leakage (Gibler et al., 2012). For instance, they transfer personal data outside the devices of end-users without their consent. In a report published by GDATA (GDATA Software, 2016), the Android platform is estimated to account for 97% of all malware on mobile devices in 2014. Over 2 million trojan applications have been detected

* *Corresponding author.*
E-mail address: aaligombe@towson.edu (A.I. Ali-Gombe).
[1] A poster version of this paper appears in CODASPY 2016 (Ali-Gombe et al., 2016).
https://doi.org/10.1016/j.cose.2017.11.006
0167-4048/

in 2015, representing a 50% increase from 2014. Modern malware is in use on an industrial scale by crime organizations and its development is often highly professional. In another report, Andrubis (Weichselbaum et al., 2014) performed an analysis on over a million (malicious and benign) apps, and found that 38.79% of the apps have data leakage. The percentage further increases from 13.45% in 2010 to 49.78% in 2014, and is also noted by Zhou and Jiang (2012). In many respects, this presents an even greater threat to users than before, as mobiles are entrusted with the most private of information and mobile malware can very effectively spy on users in real time. Overall, the security and privacy concerns surrounding these revelations increase the need for reliable and accessible app analysis systems.

Traditionally, Android apps are analyzed using either static or dynamic approaches. Static analysis involves the use of predetermined signatures and/or other semantic artifacts such as API calls, strings, etc. Enck et al. developed Kirin (Enck et al., 2009) which evaluates privacy risks based on the set of permissions requested, while Felt et al. (2011) and Zhou et al. (2012) analyzed Android applications by evaluating fine-grained API calls in addition to the permissions set. Other semantic-based analysis tools (Feng et al., 2014; Wu et al., 2012) examine components and intents in addition to the permissions and API calls made within the application binary.

Dynamic analysis on the other hand executes a target application in a contained environment (APIMonitor, 2012; Backes et al., 2013; Bartel et al., 2012; DroidBox, 2011; Enck et al., 2010; Falcone et al., 2013; Karami et al., 2013; Rastogi et al., 2013; Zhang and Yin, 2014a, 2014b). In general, static analysis has the advantage of high performance and coverage. Conversely, simple obfuscation can hinder the extraction of important data such as API names. Dynamic analysis on the other hand provides a better view of an app's behavior, although it is usually limited in scope to observed execution paths.

Most comprehensive dynamic analysis techniques either require instrumentation of the underlaying operating system code (DroidBox, 2011; Enck et al., 2010; Rastogi et al., 2013) or involve virtual machine introspection (Yan and Yin, 2012). They provide effective sandboxing for the analysis of the target applications, but unfortunately, such techniques are heavily dependent on OS versions and the Android runtime. Porting and flashing a new build on real devices for various versions of Android are not an easy task, which can limit the number and kind of applications that can be analyzed. Existing application-level techniques like those of APIMonitor (2012), Backes et al. (2013), Bartel et al. (2012), Falcone et al. (2013), and Karami et al. (2013) are mostly constrained to performing only API monitoring. Although systems like Capper (Zhang and Yin, 2014a, 2014b) can perform app-level taint analysis, their heavy reliance on static analysis for the extraction of taint slices makes it equally vulnerable to simple obfuscation.

In this paper, we present *AspectDroid*, a hybrid analysis system for Android applications based on the AspectJ instrumentation framework. *AspectDroid* performs static bytecode instrumentation at the application level, and does not require any particular support from the operating system or the Dalvik virtual machine. It weaves in monitoring code at compile time using a set of predefined security concerns, such as data/resource abuse and other non-traditional behaviors like reflective calls and native code execution. The target application is then executed on any Android platform of choice for which behavioral patterns are monitored and logged dynamically.

In summary, *AspectDroid* is a new hybrid analysis system for Android applications that has the following salient features:

Android Platform Independent: *AspectDroid* does not introduce code at the operating system level. Instrumented applications can run without any restrictions on both emulators and physical Android devices.

Adaptable to all Android Runtimes: *AspectDroid* is not restricted to the Dalvik virtual machine or Android runtime (ART).

Explicit Data Exfiltration: *AspectDroid* uses an efficient algorithm to track data propagation dynamically from source to sink.

Behavioral Tracing: We monitor applications for possible unwanted activities like telephony abuse, use of reflection, dynamic classes, and native code execution.

To determine the effectiveness and efficiency of *AspectDroid*, we carry out two different tests. In the first experiment, we analyze 105 Android apps from the DroidBench project for possible data exfiltration. The results show that the *AspectDroid* dataflow algorithm can accurately follow the propagation of target data from source to sink with 94.68% F-score accuracy. The second experiment analyzes the dynamic behavior of 100 malware samples from Drebin's dataset (Arp et al., 2014) and 100 apps downloaded from Google Play. Our findings are itemized based on data exfiltration, use of reflection, dynamic class loading, native code, and telephony abuse. The results of our analysis indicate that while phone-related data like IMEI are equally exfiltrated in both malware and the Google Play apps, that's not the case for user-related data like contacts where leak traces were more common in malware samples than the Google Play apps. Five malware samples use reflection for malicious purposes, such as invoking the methods of a background service to spoof user accounts and passwords. On the other hand, all the reflective call invocations in the Google Play samples did not result in any sensitive API call. Furthermore, we have seen more telephony abuse in malware than the Google Play apps, e.g., SMS was sent to all contacts on the phone without the user's consent. Nine malware samples invoke native processes 72 times, as compared to 6 for the Google Play apps.

We further use the malware dataset to measure the instrumentation overhead for dynamic execution. The results show that *AspectDroid* has limited memory overhead of around 1 MB and a reasonable 5.9% average CPU usage overhead.

The rest of the paper is organized as follows: Section 2 presents background on the AspectJ instrumentation framework; Section 3 provides an overview of *AspectDroid*'s design and associated algorithms; Section 4 presents the implementation of *AspectDroid*; Section 5 contains testing and evaluation of results; Section 6 enumerates some challenges and discusses limitations and future work; Section 7 reviews the related literature followed by Section 8 that concludes the paper.

2. Background

Instrumentation is the process of analyzing programs by adding trace code to their source code, binary code, or execution

environment. This provides mechanisms for an analyst to define concerns related to program verification, enforcement, monitoring, error-checking, performance, debugging, or tracing. Instrumentation techniques do not necessarily modify code but rather tamper with the execution or behavior based on defined constructs. In recent years, instrumentation techniques have gained momentum in the cybersecurity community for vulnerability (Zhang and Yin, 2014b), malware (APIMonitor, 2012; DroidBox, 2011; Enck et al., 2010; Falcone et al., 2013; Karami et al., 2013; Rastogi et al., 2013) and privacy analysis (Backes et al., 2013; Bartel et al., 2012; Jeon et al., 2012; Zhang and Yin, 2014a).

Aspect oriented programming (AOP), first introduced by Kiczales et al. (1997), is a modularized programming model allowing the separation of cross-cutting concerns (AspectJ Team, 2002–2003), which are difficult to capture in traditional programming models. AOP encapsulates the concerns, defined as `aspects`, by instrumenting extra behavior in the existing code. These aspects are special constructs forming the building blocks of AOP. Their designs can be generic, which allow for reuse throughout program execution. Implementation of AOP can be performed in two distinct ways:

1. Static instrumentation allows for code to be injected at compile time. This technique merges both the aspects and the original code into one binary, which then executes in the execution environment of the original code.
2. Dynamic instrumentation, on the other hand, injects code at runtime. In most instances, this requires a custom classloader to enable the interpreter to understand and implement the AOP features.

In 2001 (Kiczales et al., 2002), PARC developed an extension for AOP designed specifically for the Java programming language, called AspectJ. Its Java-like syntax, coupled with its ease of use, makes AspectJ a very popular instrumentation tool for Java programs. Aspects in AOP are defined by some key terms:

1. Pointcuts are defined by `kinded` constructs such as function `call`, method `execution`, `within` class, `cflow`, etc., which match some specified `signatures`.
2. Signatures are semantic definitions which can be decoded by the AspectJ compiler during joinpoint creation. It can encapsulate both broad and narrow definitions, giving an analyst ample flexibility.
3. Joinpoints are points within the execution of the program that are interesting and/or defined by the concerns of an analyst. These are chosen based on constructs defined in a pointcut.
4. Advice is the piece of code that gets executed when a certain joinpoint is reached during program execution

In addressing security concerns, advice defined for a joinpoint adds some functionalities such as logging, code injection, value manipulations, execution rerouting, skipping execution, etc. to an instrumented program. Advice to be executed can target `before`, `after`, and `around` the execution of a particular joinpoint. As the name implies, execution of `before` advice precedes the execution of the target joinpoint. In this advice, parameters and the target object can be retrieved, in addition to signatures, source location, etc. For `after` advice, in addition to the information extracted in before advice, the return value can also be retrieved and evaluated. The most interesting is `around` advice which, although potentially expensive to use, allows code injection and modification of arguments, variable values, and return values.

The code snippet in Listing 1 shows a sample aspect that defines a pointcut which picks a joinpoint at the call to `getDeviceId`. When instrumented, this aspect will pick the method `getDeviceId` from the class `TelephonyManager`. The joinpoint is picked because the signature matches the method in that class and it is the only class in the Android SDK with such a method. However, if within the application there also exists a library class with such a method, our broad signature will automatically capture such a joinpoint, too.

Listing 1: Simple Aspect

```
public aspect Logger{
  pointcut myId():
     call(* *..*.getDeviceId());
  after()returning(String id): myId(){
     log.v("AspectJ", "DeviceId="+id);
  }
}
```

2.1. *Bytecode weaving*

In the Java compilation process, an intermediate representation called `bytecode` is generated when the original source code is compiled. This bytecode is contained in class files representing each source class. More specific to Android, the system has added another level of abstraction to its compilation process, where the class files are further compressed into one dex file. Bytecode weavers are tasked with weaving together class files (both Java classes and aspect classes). In this paper, our chosen bytecode weaver is AspectJ (Kiczales et al., 1997). Its robust framework allows us to define and inject security concerns related to Android apps for the purpose of logging and monitoring. More so, its programming syntax and semantics are identical to those of the Java language, allowing us to tie and weave the monitoring code into a target Android application with better precision.

AspectJ compilers (ajc) can accept both raw sources and class files for compile-time weaving and thus have the capability to compile and weave the aspect/Java sources and/or class files to produce a new woven class. The resulting merged Java bytecode has to be compatible with the execution platform's VM. However, in load-time weaving AspectJ exposes an interface that facilitates the weaving process between the target bytecode and a custom `classloader`.

3. System design

AspectDroid is a hybrid system that uses static instrumentation to inject monitoring code into the target app based on some specific cross-cutting concerns. Requirements of our system involve analyzing unknown binaries where there is no available source code. The core of *AspectDroid* is built based on

compile-time bytecode weaving. This form of static instrumentation takes the advice defined in an aspect and weaves them at specified joinpoints in a target class file. For Android apps, the resulting binary is `dexed` and repackaged into a new apk. Since this new application does not need a custom `classloader`, it has the flexibility of executing on any device emulator without changes to the underlying OS.

With *AspectDroid*, the new injected code executes alongside the original code and performs custom logging and other analytical functions. The instrumentation engine (IE) which is the primary component, forms the backbone of *AspectDroid* and is designed to address three objectives:

1. Dataflow Analysis
2. Resource Abuse Tracing
3. Analytics of Suspicious Behavior

Our instrumentation code is encapsulated in an aspect and is tailored for each of the objectives mentioned above. The aspects are weaved into the target app using AspectJ's ajc compiler, producing the instrumented version used to perform the analysis. The instrumentation process is done in-vitro on a host machine. After successful recompilation, the target app is then pushed onto the test bed for execution (Fig. 2).

3.1. *Dataflow analysis*

AspectDroid performs application-level tainting of target data source(s). Our approach is built around the fact that standard Java and Android libraries use specific method naming conventions to express common types of operations. Thus, we utilize the consistent use of specific verbs, such as `read`, `open`, `write`, `put`, `connect`, and `execute`, to define broad signatures to capture actions such as file/stream/network access. More specific signatures, such as `getLongitude`, are used to define narrower joinpoints. Based on all the signatures, we define pointcuts to select various `source`, `sink`, and `propagation` joinpoints. With the help of AspectJ APIs, a joinpoint's data, such as the `target object`, `parameters`, `return values`, etc. (as shown in Fig. 1) can be extracted at runtime. Java programming semantics categorize data types as primitive, object, and arrays. Although beyond the scope of this paper, it is important to note that the JVM stores and processes these data types very differently. Therefore, our data sources, propagation and sink for each data type are handled differently. To simplify the terminology, we refer to `primitive` data types, such as `string`, and `character`, as `low-level data types`.

Our dataflow analysis is limited to explicit propagation, where the tainted data must be in the sink call, as shown in Listing 2. On the other hand, Listing 3 illustrates an example of an implicit flow which exfiltrates inferred data based on the real tainted data.

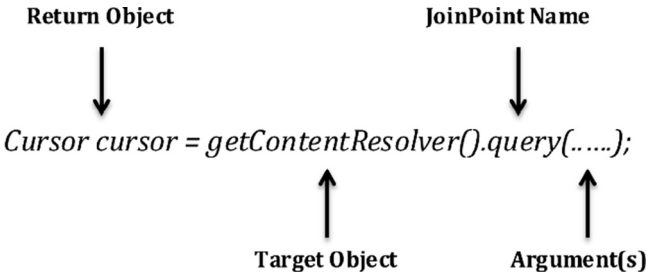

Fig. 1 – Parts of a method joinpoint.

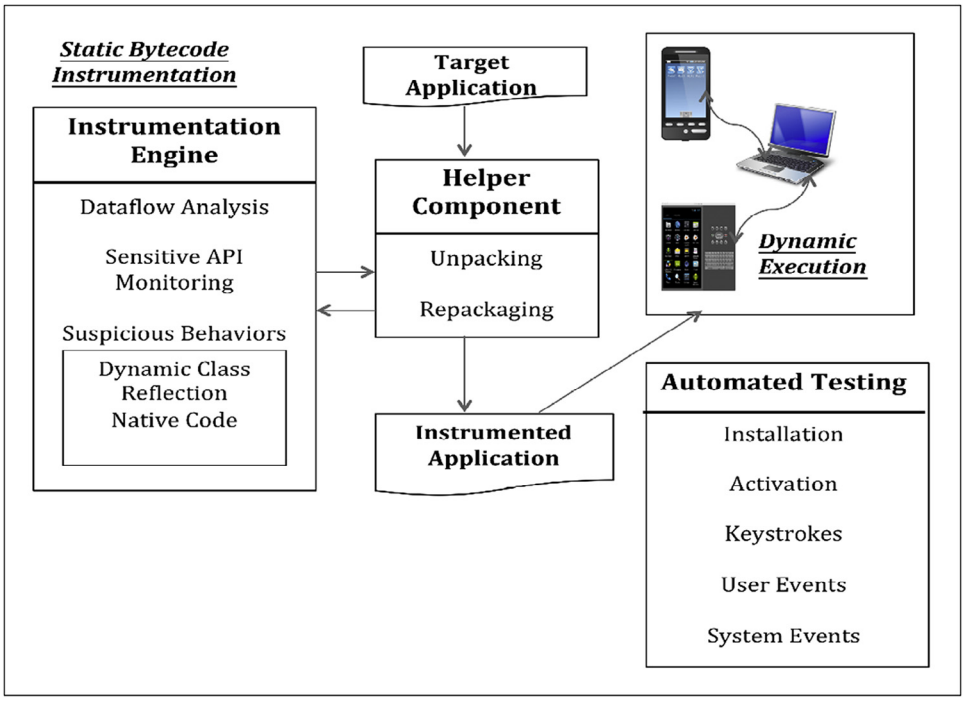

Fig. 2 – *AspectDroid* implementation architecture.

Listing 2: Explicit Data Flow

```
TelephonyManager telephonyManager =
    (TelephonyManager)
   getSystemService(Context.TELEPHONY_SERVICE);
String IMEI = telephonyManager.getDeviceId();
if (!IMEI.equals("00000000"){
   String id =
       Base64.encodeToString(IMEI.getBytes(), 0);
   SmsManager sms = SmsManager.getDefault();
   sms.sendTextMessage("5556", null,id, null, null);
   }
```

Listing 3: Implicit Data Flow

```
TelephonyManager telephonyManager =
    (TelephonyManager)
   getSystemService(Context.TELEPHONY_SERVICE);
String IMEI = telephonyManager.getDeviceId();
if (!IMEI.equals("00000000"){
   String val = "Device not emulator";
   SmsManager sms = SmsManager.getDefault();
   sms.sendTextMessage("5556", null,val, null, null);
   }
```

The AspectJ API used by our system is not designed to create joinpoints on conditional/branch instructions. Nonetheless, in the analysis of Android applications, sensitive data leaving the device is the real threat, not inferred data. As illustrated in Listing 2, the real device IMEI was exfiltrated compared to "Device not Emulator" in Listing 3.

3.1.1. *Taint sources*

We are interested in sources that are relevant to the privacy and security of the user. We define `vital` sources as phone-related data, content provider objects, file reads, and user input. In Android, most important data are guarded by permissions and only accessible to the user through specialized Android API calls. Other relevant data not guarded by permission, such as data read from files and user input from textboxes, are also accessed via the standard Java/Android APIs. Specialized pointcuts are created using signatures to intercept these vital API calls. After execution, the return value is stored as a key in a `taint map` with a corresponding special tag for each unique source as the value. Depending on the return type, low-level data types are stored in raw form, while every other object is stored in hash form. This storage design is very significant in reducing the overhead associated with checking if tainted data is part of an object. It allows us to check if the object is tainted using its hashcode at propagation or sink joinpoints.

3.1.2. *Taint sinks*

Taint sinks are defined as points where the target application communicates with an external component, either within the device or the outside world. In our dataflow analysis, we seek to monitor only those sinks that form a `possible` exfiltration point for the data sources defined above. The data sinks are broadly categorized as network, e.g., writing to a `Socket`, `URLConnection`, etc.; SMS sends; file writes (both ordinary files and shared preferences); and IPC. We use the same signature semantics to pick the sink joinpoints. We also leverage the `around advice` of such joinpoints to check if its arguments, or target, contain tainted data.

This process is straightforward for parameters. For example; if the sink call is a `sendTextMessage`(..), the tainted data will be checked against the parameters of this joinpoint. However, for a target object we need to parse it and check the associated fields against the keys in the taint map. For example; if the tainted data is appended to a URL, and then a URLConnection is created from that URL object, which then invokes it's `getOutputStream`() method. Our system will have to parse the URLConnection object to get the URL field and compare that against the data in our taint map. Overall, data exfiltration is detected if a tainted piece of data is found either within the sink joinpoint's parameters/parameter's fields or within the target/target's fields.

3.1.3. *Taint propagation*

Knowing data sources and sinks alone cannot accurately determine data exfiltration; we also need to identify the data propagation process as represented by the sequence of variable assignments along the path from source to sink. The tainted data can be part of an object's field and the object can be manipulated in different ways. For every joinpoint, we check if it contains tainted data; if so, the appropriate propagation rule is picked based on the respective joinpoint's return type as enumerated in the 7 point rules below:

1. Rule 1: Joinpoint that returns a low-level data type and contains a tainted argument.
2. Rule 2: Joinpoint that returns a low-level data type and contains a tainted target.
3. Rule 3: Joinpoints that create an array from other tainted data types.
4. Rule 4: Joinpoints that convert a tainted array to other data types.
5. Rule 5: Object constructor joinpoint that contains a tainted argument.
6. Rule 6: All joinpoints with object return type that contain tainted arguments.
7. Rule 7: All joinpoints with object return type that contain tainted target.

For joinpoints targeting low-level data types, their return values and target object are the same. However, for object joinpoints, the target is always a reference to the location of the object in memory while the return type could be anything. For example, an object's joinpoint could return a Boolean indicating success of a method call, void, a low-level data type, or other objects. This distinction forms the basis of how our taint tag/map is updated after the execution of the joinpoint. Table 1 gives a taint propagation example for each of the flow rules, and shows the taint tag/map update after the joinpoint's execution. Propagation rule 7 can create a weaving process that might get out of hand, thus we included some optimizations for joinpoints associated with that rule, based on the object's class.

To optimize the weaving process and reduce the complexity of the instrumentation, the propagation's joinpoints for every source are created along the control flow path of its enclosing

Table 1 – Flow rules examples for updating taint/tag map.

Rules	Joinpoint Example	Taint Data	Taint Tag/Map Update
Rule 1	Int myInt = System.identityHashCode(val)	valueOf(val), tag = DeviceID	valueOf(myInt), tag = DeviceID
Rule 2	String str1 = myInt.toString()	valueOf(myInt), tag = DeviceId	valueOf(str1), tag = DeviceID
Rule 3	char arr[] = str1.toCharArray()	valueOf(str1), tag = DeviceID	Hashcode(arr), tag = DeviceID Elements of arr, tag = DeviceID
Rule 4	Str str2=Arrays.toString(arr)	Hashcode(arr), tag = DeviceID	valueOf(str2), tag = DeviceID
Rule 5	StringBuilder stb = new StringBuilder(str2)	valueOf(str2), tag = DeviceID	Hashcode(stb), tag = DeviceID
Rule 6	stb.append(val2)	Hashcode(stb), tag = DeviceID valueOf(val2) tag = LineNum	HashCode(stb) tag = DeviceID and LineNum
Rule 7	Vector vec = new Vector() vec.add(str2)	New empty vector is created valueOf(str2), tag = DeviceID	Hashcode(vec), tag = DeviceID

method. For example, if the data source IMEI returned by `getDeviceId` is found within the body of an activity's `onCreate` method, then the propagation joinpoints will only be created for methods that satisfy the propagation rules above and are in that control flow. This optimization greatly enhances our weaving process and eliminates the need for redundant joinpoints.

3.2. *Resource abuse tracing*

Access to some vital functionalities such as Telephony (SMS and Calls) on the mobile devices are requested through specialized API calls. According to a 2012 Trend Micro report (Trend Micro, 2012), resource abuse is the most common category of Android malware. Thus it is imperative for an analysis system to trace and report such abuse. With *AspectDroid*, the system instruments the telephony methods invocations and have their target object, parameters, and return value logged. This information is significant in determining the phone number used (premium service or device contact), the message content, settings, and format.

3.3. *Analytics of suspicious behaviors*

Programming practices such as native call invocation, dynamic class loading, native code execution and reflective call invocation add flexibility to software development. Although these concepts may be benign, malware can often hide its behavior using these practices to hinder static analysis or for malicious purposes such as privilege escalation.

Reflection, for instance, allows method calls to be resolved dynamically at runtime. Malware can use this technique to hide calls to sensitive APIs. With *AspectDroid*, we instrument reflective calls and analyze the target object at runtime for possible tainted data sources, propagation, sinks or any sensitive API calls. We also check their parameter arrays for possible taint propagation. For suspicious behaviors other than data exfiltration we instrument the *"method.invoke"* function call to log the declaring class of the target object which resolves to the method name and class. We also record the arguments which contain the instance of the class object and the parameter array needed to invoke the target method. Listing 4 shows the code snippet:

Listing 4: Pointcut for Reflection Call Invocation

```
pointcut getReflect(): !within(Logger) && call(*
    java.lang.reflect.Method.invoke(..));

Object around(): getReflect(){
   Object [] params = thisJoinPoint.getArgs();
   Object tar = thisJoinPoint.getTarget();
   logVals(tar, params); //Convert the objects to
       human readable form and log
}
```

For malware analysis, the *logVals* routine is designed to check for suspicious method invoked using reflection such as *getInstalledPackages*.

Dynamic class loading is another programming concept in Android that allows app to load extra classes at runtime. One of the drawbacks of static bytecode instrumentation (and by extension all static analysis) is that only available classes are processed at compile time; extra classes loaded at runtime are not affected by the weaving. To address this, *AspectDroid* implements `dynamic class instrumentation`: at the joinpoint where `Dexclassloader` loads the new dex file, the weaved advice captures the absolute path to the file, sends it to the host machine via an Asynchronous task, and waits for notification to proceed. On the host machine, *AspectDroid* has a server side component that receives the dynamic class, instruments it and pushes it back to its original path on the testbed. On return of this an Async task, normal program flow resumes. Although this wait time slows down the process, it considerably expands the code coverage of our analysis.

AspectDroid also logs native code invocation, both for simple processes like Logcat or through the Java native interface. Although it does not trace the activities within the native code, it does log the name, object, parameters, and return value. This logging not only enables us to view what native functions are invoked by a target app but can also allow us to know the location of the code. This functionality is especially significant for a native payload that is downloaded dynamically at runtime. While our system is not designed to parse the native code, however with some manual effort analysts can be able to pull the executable file and reverse engineer to get more information.

4. Implementation

4.1. *Prototype implementation*

We implemented a working prototype of our system in Python, Java, and PHP. The instrumentation engine is set up on a host machine (64-bit Ubuntu system) for the initial dex weaving and dynamic class instrumentation. Our software dependencies include external tools and libraries; `dex2jar` (Bob, 2014), `AspectJ-ajc` (AspectJ Team, 2015), `Apache Web Server` (Apache Software Foundation, 2013), `Apache Commons` (Apache Software Foundation, 2015), `aspectjweaver` (AspectJ Team, 2015) and `Android SDK` (Android-Studio, 2015). Our initial experiments were carried out on a physical device (a rooted Motorola Droid2 with Android 2.2) and two emulated devices (Android 4.1.2 and 4.4.2). The execution environments are loaded with text messages, calls, contacts, one Gmail account, and some browser history. To buttress our claim that *AspectDroid* is environment-agnostic, we performed a second round of testing on HTC-One S9 running Android 6.0.

4.1.1. *Helper component*

AspectDroid includes a "helper" component containing modules that automate key actions, including unpacking, repackaging, and application signing. Android applications are written in Java and compiled into a compressed class called `classes.dex`. However, the AspectJ compiler does not understand the dex file format, thus the need for decompression before weaving. We use a popular open source tool called `dex2jar`, which takes an application file (.apk) or classes.dex as input and outputs a jar file containing individual .class files. When the target application is unpacked, it can be weaved together with desired aspects. After the instrumentation process, the class files are repackaged (dexed and zipped) and resigned into an Android-compatible app using `jar2dex` and `versign` respectively.

4.1.2. *Automated testing*

Unlike many traditional applications, smartphone apps are mostly event-driven and exhibit their true functionalities based on user interactions and in response to system events. For example, forcing an SMS to be received so that a broadcast receiver can be activated is an important system event that needs to be triggered for us to observe SMS abuse.

In the case of bulk analysis, manual execution of apps and triggering such events can be time-consuming. One of the drawbacks of dynamic analysis is code coverage and a single execution path corresponding to a single app execution, whereby information obtained may not necessarily represent the complete behavior of the target app. Our assumption is the more a tool can explore an app, the more information about the app's behavior can be obtained. For that reason, we build into *AspectDroid* an automated testing module as Python scripts which triggers a series of system and user events to more fully exercise an app's functionality. This module combines some open source tools together with custom-built instrumentation programs. These events are designed to mirror real-life events on a regular Android device. They include:

1. App installation and activation of its main activity, as specified in the manifest, using `adb`.
2. Random keystrokes that simulate user touch and gestures on the app using `monkey`.
3. A user input is simulated where necessary within the instrumentation framework. `EditText` user inputs can be associated with different input types. Most developers specify input types as provided by the Android API – email, password, etc. We make a best effort to generate data to match its possible input type. This program is attached to the body of the instrumentation code.
4. SMS, calls and device settings are generated and manipulated using uiautomator while GPS coordinates are simulated and triggered on the emulator by `telnet`.

Independent testing frameworks like Android Monkey are limited to only random application touches and gestures. With our automated testing, the simulated user input built on the `EditText-SetText` method automatically creates the needed `textbox` data during analysis, which proves to be very important. For example, if an EditText is expecting an email, if the Ok button is hit using Monkey, the application may return an error and program execution may not proceed due to an empty text. But with our injected input text, execution will proceed without an error.

Other vital parts of this testing module built with uiautomator help with forcing various system's event like `Calls`, which would otherwise have to be done manually.

5. Testing and evaluation

Our approach seeks to provide analysts with an easier to use and more flexible system for application analysis. It is capable of examining and monitoring Android applications without restriction based on version and/or platform while still maintaining a very high level of accuracy. The objectives of the evaluation were to quantify the following aspects of the system's performance:

1. *Accuracy*. We tested the accuracy of our dataflow algorithm on 105 applications from the DroidBench corpus.
2. *App Analysis*. We further evaluate the effectiveness of our system by comparing the behavioral patterns in 100 real malware families from the Drebin dataset and a set of 100

apps downloaded from Google Play. We examine data exfiltration, telephony abuse, reflective invocation, dynamic class loading, and native code execution.
3. *Execution overhead*. We measured the cost associated with dynamic execution of the target app post-instrumentation.

5.1. Accuracy of dataflow algorithm

DroidBench 2.0 (DROIDBENCH, 2015) is an open source project consisting of 120 simulated Android applications used for testing analysis tools. These applications evaluate the accuracy of an algorithm in detecting dataflow between a source and a sink. The authors employ different methods of data manipulation, such as callbacks, arrays, application lifecycle, inter-application communication, loops, reflection, threading and implicit flows to hide the flow of sensitive data. The apps are relatively small and they may not necessarily be representative of real-life apps and/or malware in terms of size. However, they contain a wide spectrum of diverse, tricky dataflow paths that can be employed by malicious and/or over-privileged applications, thus making them a corpus of interest to test *AspectDroid*.

Before executing the apps with *AspectDroid*, we execute the untampered dataset to determine if they are running correctly and producing expected results. Out of 120, 15 apps failed to execute correctly in our environment due to either permission errors or other bugs and were excluded from the analysis. The remaining 105 apps were instrumented using our *AspectDroid* prototype.

Based on the original source code for the 105 apps, the ground truth indicates 86 apps have data leaks and 19 apps have no leaks. Our experiments show that *AspectDroid* yielded 80 true positive (TP) results, 16 true negatives, 3 false positives, and 6 false negatives. Thus, the *AspectDroid*'s precision is 96.4%, recall is 93.02%, and the standard F-measure stands at 94.68%. Subsequent analysis showed that in the three false negative cases, tainted and untainted data were added to a data structure, then the app sinks only the untainted data. Our algorithm taints an object that contains a tainted field, entry or element and does not handle removal of that data/object from taint map once it is written. Since the untainted data is still part of a tainted object, we recorded a false positive. With respect to the six false positives, four were apps with the following propagation paths: `Public API Field1`, `StartProcessWithSecret` and `Implicit Flows`. Our dataflow algorithm taints by means of data comparison (possible taint with items on taint map), thus data exfiltration that is not explicit cannot be detected. The remaining two under tainting were a result of an optimization added to our propagation rule in order to reduce the effect of over-weaving (which results in too much additional code added to the application). This optimization is a tradeoff between the effect of over-weaving and a possible false negative; hence, these two false results are avoidable.

5.2. App analysis

To test the effectiveness of *AspectDroid* for analyzing Android applications for violations of security and privacy concerns, we used malware samples from Drebin (Arp et al., 2014) dataset, a corpus comprising 179 malware families. In our experiments, we picked one sample per family from the top 100 families. For the non-malicious samples, we downloaded 100 Android apps from Google Play. All 200 samples are instrumented, recompiled, and executed using our automated testing module. In our prototype we tagged 27 important data sources, including phone-related data (IMEI, IMSI, ICCID, line number, and location data), database queries, and input data. We also created joinpoints on some sensitive APIs that perform telephony functions, native code execution, dynamic class loading, and reflection invocation. The dataflow and sensitive API traces created after each app execution are then parsed using a Python script to obtain the aggregated result. We categorize the analysis result into 4 groups: data exfiltration, telephony abuse, reflection and dynamic class loading, and native code execution.

5.2.1. Data exfiltration

Malware and to a large extent privacy-agnostic applications often target user and/or phone-related data either with malicious intent, for advertisement or identification/records purposes. Most sensitive data are guarded by one-time permissions (for Android versions 1–5) that give an app open access to quite a large group of data on a device e.g., Phone-state permission. We define exfiltration as unauthorized writes of sensitive data to a file (log, sharedprefs, user-defined files), the network, or SMS that are not explicitly granted by the user at the point of transfer. Our analysis of the 100 malware samples showed 127 explicit data exfiltration paths of the 27 tainted sources carried out by 23 samples. Our results showed IMEI, IMSI, ICCID and LN are the most widely exfiltrated phone data. This is followed by contacts, call logs, and SMS from user-related data. SharedPref and Network are the most common sink calls we observed while SMS was the least observed. For the Google Play apps, we observed 25 exfiltration paths, most of which are location and phone IMEI. Network is the sink path for all these data leaks.

5.2.2. Telephony abuse

SMS is one of the most widely abused resources on Android smartphones. Out of the 100 samples we evaluated, 8 samples were recorded to have some level of SMS abuse. The Pirater malware sends SMS to all contacts on the user's phone, posing as the user. The socially engineered, "friendly" SMS generated by Pirater contains a link that downloads the same malware to the receiver's phone if clicked. The MobileTX malware, on the other hand, does not just abuse SMS functionality, but also transmits the phone's ICCID to a private number via SMS. The remaining 6 samples send specially crafted SMS to premium numbers. We have not recorded any phone call interceptions, spoofing or recording in any of the analyzed malware. We observed the use of SMS in 2 apps and Calls from 3 apps which belong to the communication category on Google Play. In all these instances, the SMS and CALLs were authorized by the user, based on user-supplied input.

5.2.3. Reflection and dynamic class loading

The reflection API is part of the standard Java environment and allows method calls to be resolved dynamically at runtime. It is a powerful tool that can be employed by malware to evade static detection. We have observed 5 malware samples that use

reflection in different ways. We then examine if such invocation exhibits some element of malicious intent. The Mobsquz and FakeDoc malware reflectively check if the device has support for telephony-related services (phone calls and SMS). Although this may not necessarily constitute malicious behavior, given the functionality of the applications as an antivirus scanner and battery optimizer, it requires further analysis. The FaceNiff malware uses reflection to invoke the methods of a background service that spoofs user accounts and passwords after it has successfully executed the super user command. The 2 other remaining families, BaseBridge and DroidDream, are not suspicious as they both invoke methods from GUI-related classes. We observed 34 instances of reflective call invocation on the Google Play apps. Surprisingly, this is higher than reflective call invocations recorded within the malware samples. We also observed that the BaseBridge malware dynamically creates 3 jar files (bootablemodule.jar, moduleconfig.jar, mainmodule.jar) and 2 dex files (mainmodule.dex and bootablemodule.dex). Within the timeframe for our automated testing and even with an extended manual execution afterwards, the app did not load these new classes dynamically as expected. Thus, we rewrote the binary to force the app to load the new dex files. This enables our dynamic instrumentation to trace the loading joinpoint and the newer classes were instrumented using our dynamic instrumentation engine. For the Google Play apps, 7 dynamic classes were loaded in 5 apps within our testing time. We were able to successfully instrument and execute all the dynamic classes.

5.2.4. *Native processes*

Android applications are commonly written in pure Java code, although quite a number of them include an embedded C/C++ binary. Over the years, Android malware has exploited this capability to embed mostly root exploits that trigger privilege escalation. In other instances, Linux commands that communicate with the underlying Android kernel are becoming increasingly common. In our dataset, 9 out of the 100 malware samples invoke native processes 72 times. Commands like `su`, `chmod`, `ps`, `mount`, and Android's `logcat` are the most widely executed native processes. We have also noted the execution of an unknown binary (myicon) in DroidKungFu malware. Since *AspectDroid* does not instrument native code, we log the code path and then manually extract the code using adb. An Md5Sum later verified that the native binary is a root exploit belonging to the family `RageAgainstTheCage`. We've noticed native code execution in 6 out of the 100 Google Play apps. In comparison with the malware apps, the Google Play apps all executed ".so" libraries vs. starting other processes like `chmod` or `su`. Beyond noting that a particular native code has been called within the Java execution, *AspectDroid* does not monitor the native code's behavior, as the instrumentation engine works only on Java. Thus it is inconclusive what some unknown native libraries do.

5.3. *Runtime overhead*

The most important costs of instrumentation occur at runtime, since both CPU and memory usage are vital on a resource constrained machine. It is especially important that apps limit their resource usage to avoid possible garbage collection. Though uncommon in foreground processes, this does occur when apps consume too many resources.

The CPU usage is the percentage of CPU time used by a process. We measured the value given the system uptime (uTime), process start time (startTime), and the CPU time spent in both user and kernel code for the main process and any of its child processes (uTime, sTime, cuTime, csTime). The formula is given below:

$$\begin{aligned} seconds &= upTime - (startTime/Hertz) \\ tTime &= uTime + sTime + CuTime + csTime \\ cpuUsage &= ((tTime/Hertz)/seconds) * 100 \end{aligned} \tag{1}$$

We carried out this experiment by re-running the 100 malware families using automated testing on the same platform, keystroke seeds, and number/pattern of system and user events. Using the `procrank` utility, we obtained the process memory size from each app both before and after instrumentation as well as the CPU indices above. The experiment was executed 5 times and an average for each metric (Memory and CPU) was computed.

The dark portion of the stacked bar chart illustrated in Fig. 3 shows the memory usage for each malware pre-instrumentation, while the lighter shade shows the overhead after instrumentation. The data illustrates that the `MemSize` difference is uniform and on average, 1 MB of additional memory is required to execute the instrumented application. This translates to *approx* 16% more memory usage on average. In our tests, this overhead caused no issues with any of the apps.

Fig. 4 on the other hand shows the percentage of CPU needed to render and execute each malware. The dark portion indicates the CPU usage before instrumentation while the lighter portion stacked showed the CPU usage overhead. Although the results are not uniform, the average CPU overhead is approximately 5.91%. Some apps tend to have significantly more overhead than others. We manually examined these apps and found two important factors: the number of data sources tagged and the propagation path can have varied and compounded impact on the CPU usage overhead. CPU intensive apps like games that request a lot of tagged data, and especially if these requests are along the path of an activity, tend to require more CPU time to load the activity (e.g., the PJApps and Jifake malware). Although the Fujacks malware has the highest CPU usage pre-instrumentation, its overhead is negligible since it did not request any tagged data.

6. Challenges and discussion

In the evaluation section we discussed the accuracy of the *AspectDroid* algorithm in detecting data leaks, the importance of tracing resource abuse and detection of suspicious behaviors like reflection, native code and dynamic class loading. Furthermore, we also highlighted the overhead associated with our system. Naturally, some challenges remain. In bytecode weaving, the compiler has to make a best effort adjustment

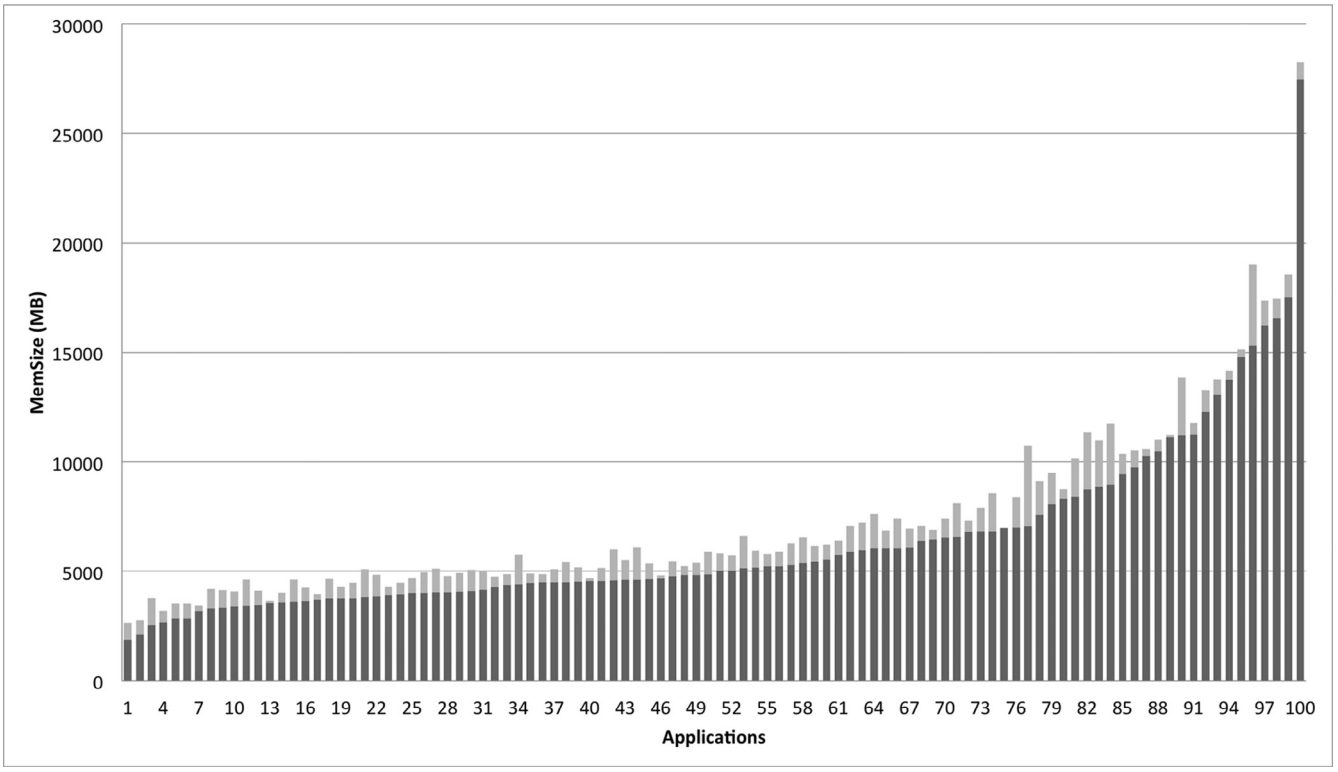

Fig. 3 – MemSize overhead (MB).

to registers, fields, methods and instructions of the weaved class. Some applications can be sensitive to this kind of intrusion and as such can pose a setback in our recompilation process. We make a best effort to optimize the weaving process, especially in dataflow aspects, while at the same time keeping false negatives as low as possible. Specifically, we ensure that:

1. Propagation rule 1, which handles primitive returns, excludes void and Boolean values. Allowing Boolean values in our taint map significantly increases false positives.
2. GUI-related classes that handle graphics, views, and activities are also excluded from propagation in propagation rule 7.

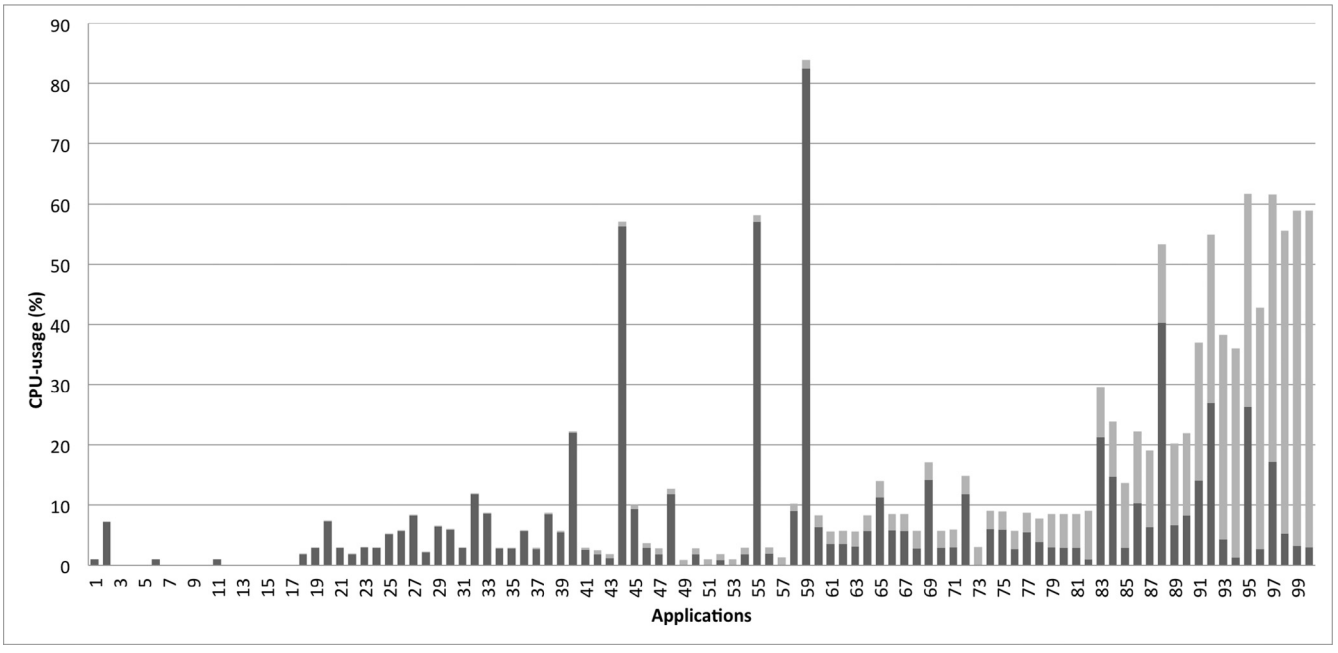

Fig. 4 – CPU usage overhead (%).

3. Well-known public libraries from Android, Amazon, Google, Samsung, and Apache are excluded from the scope of weaving, however their calls are included within the signatures, if necessary.
4. We use abstract methods within advices to reduce the number of instructions added directly to the weaved class.

Overall, our system effectively uses these optimization techniques to boost its accuracy and performance. In our testing, *AspectDroid* has proven to be effective in analyzing dataflow paths, sensitive API monitoring and analysis of suspicious behaviors. It provides a flexible and efficient system for assessing Android applications and does so with relatively low overhead.

6.1. Limitations

The main limitations of every static bytecode instrumentation are anti-unpacking and anti-repackaging obfuscation mechanisms. Developers can include obfuscated bytecode in their compiled dex files that decompilers cannot parse correctly. However, in most cases this obfuscation does not affect method invocation, which is what *AspectDroid* uses to create joinpoints. Also, malware can detect instrumentation code and/or change the package signatures, which can negatively affect analysis with *AspectDroid*.

The AspectJ instrumentation framework is limited to processing only method instructions. As opposed to variable level tainting, *AspectDroid*'s dataflow compares the hashes of raw data and as such cannot be affected by simple manipulations through variable re-assignment. However, arithmetic instructions can have an adverse effect on our taint propagation (though we have not encountered such in our analysis). As mentioned in Section 3, joinpoints on conditional instructions cannot be created due to limitation in the AspectJ's APIs. This limits our dataflow analysis to explicit data exfiltration.

Another limitation of our approach is analysis of native code. At this point, *AspectDroid* can only trace to the point where a native class is loaded and executed and it can return the name and parameters for the execution. However, it cannot trace inside native code. Very few Android applications use native code and even with malware, the native code is typically used only for privilege escalation which is heavily dependent on system vulnerabilities.

6.2. Future work

As part of future work, we are working on improving our automated testing module such that all control flow paths are forced to execute. Using code refactoring, we intend to inject simple methods after arithmetic and conditional instructions that analyze the preceding instructions' parameters. Then joinpoints will be created for the new method call at weaving time. This will take care of possible mis-propagation, thereby improving our dataflow analysis. Furthermore, the ability to analyze native code will significantly improve the scope of *AspectDroid* and as such, we intend to include a debugging architecture in a future revision of *AspectDroid*. Within the instrumented advice during native code execution, a debugger can be started with the ID of the new process to collect lower level syscalls made by the native code.

7. Related work

7.1. Application-level instrumentation

The first application-level dynamic taint tracking on Android was developed by (Zhang and Yin, 2014a, 2014b). Their system, called Capper, is designed to monitor exfiltration of sensitive data from source to sink. However, their work requires a large amount of static analysis to refactor the Java bytecode and compute taint slices, which are used at runtime as the taint propagation map. This system is prone to most of the inaccuracies of static taint tracking that can result from simple obfuscation techniques. Furthermore, data sources, propagation, and sinks that pass through reflective call invocation are not processed if the invoked class and method names cannot be statically resolved. And finally, like most app-level analysis systems, Capper does not handle dynamic class instrumentation. Our system, on the other hand, can perform better dataflow analysis since it can handle Java reflection and runtime class instrumentation.

Another research effort that target app-level instrumentation is APIMonitor (2012). The authors of DroidBox developed APIMonitor to counter its numerous porting issues. Afonso et al. (2015) adopted this technique to detect Android malware by mining dynamic features (API and system calls). Furthermore, the work of Backes et al. (2013), Bartel et al. (2012), Falcone et al., 2013 (RV-Droid) and Karami et al. (2013) all leverage static bytecode instrumentation to analyze method calls in target applications at runtime. Although they use different instrumentation frameworks, these systems are all limited to sensitive API monitoring during program execution. In contrast, *AspectDroid* is a complete analysis system that targets security concerns such as dataflow analysis, sensitive API monitoring, as well as analytics of suspicious behaviors.

7.2. Low-level instrumentation

Most Android dynamic analysis tools are developed by instrumenting the operating system code and/or the underlying framework. TaintDroid (Enck et al., 2010) is a real time dynamic taint tracking system that monitors the flow of sensitive data. It uses some basic dataflow rules to track the movement of tainted variables, method files, and IPC messages from sources until they reach a specified Java library sink.

Several extensions to TaintDroid (DroidBox, 2011; Rastogi et al., 2013; Weichselbaum et al., 2014) were built with added functionalities. DroidBox (2011), for instance, logs an app's activities related to starting services, broadcast receivers, SMS, and calls made, cryptography operations performed using the Android API, and file read/write operations, irrespective of taint marking. Andrubis (Weichselbaum et al., 2014) is an automated analysis system that combines both static and dynamic approaches to an app's analysis. Applications submitted via an online link are dynamically examined on a QEMU-based emulation environment for method tracking, system level analysis and data exfiltration using TaintDroid. Other systems like AppsPlayground (Rastogi et al., 2013) added more functionality, such as kernel level-monitoring and automated testing, to TaintDroid. These approaches rely on low-level instrumentation,

making them very OS version-dependent and in some cases platform-dependent. Importantly, TaintDroid-based systems are very dependent on the Dalvik virtual machine and as such will require a complete make-over to port to the new Android runtime. Furthermore, stealthy malware can often detect emulation environments which may result in inaccurate analysis (Mutti et al., 2015). Lastly, due to significant requirements for expert knowledge to port from version to version, the capacities of such system for long term analysis is very limited.

Other host-based dynamic analysis tools are CopperDroid, DroidScope, AASandbox and Crowdroid. CopperDroid (Tam et al., 2015) is virtual machine introspection (VMI) technique that reconstructs malware behaviors by observing and dissecting system calls which result in the identification of interesting OS and high-level Android-specific traits. Another technique that uses VMI is DroidScope (Yan and Yin, 2012). This system monitors the activity of untrusted applications through API tracing, native instruction and Dalvik tracing, and taint tracking. AASanbox (Blasing et al., 2010), on the other hand, evaluates system call logs by placing hooks between kernel space and user space. These hooks hijack the system calls made and log information such as process ID, syscall name, and execution time. CrowDroid (Burguera et al., 2011) analyzes system calls performed by an application based on logs collected using the strace debugging utility in a lightweight CrowdClient. This system is limited to extracting only Linux-specific information like open files, but cannot give broad information on IPC and Android-specific data. Afonso et al. (2016) perform runtime analysis of the Android native code in a sandbox to augment static analysis techniques that operate only on the Java code. Most of these sandboxing techniques are built on a single Android version, thus restricting its use to that particular release. If tasked to analyze malware not developed for such release, the tool may likely fail. While for *AspectDroid*, it is entirely an app-level analysis system which cannot be affected by new Android release either in SDK version or runtime.

More recently, dynamic binary instrumentation (DBI) systems like DynamicRIO (Bruening et al., 2012), PIN (Luk et al., 2005), Spike (Vasudevan and Yerraballi, 2006), and Dyninst (Buck and Hollingsworth, 2000) that perform runtime monitoring are very dependent on low-level system operation. With the exception of PIN, these are not applicable to ARM systems. DBI techniques are also very dependent on the underlying hardware architecture and as such will require modification of the operating system to perform Android app analysis.

8. Conclusion

In this paper we have discussed *AspectDroid*, a hybrid system for Android app analysis, which provides an efficient and flexible alternative for detecting suspicious and illicit behavior independent of Android runtime and or system releases. Our goal is to ease analysis and avoid the numerous problems associated with porting between versions and building a customized device kernel.

The instrumentation engine which is at the heart of *AspectDroid* is designed to achieve three main objectives: dataflow analysis, detection of resource abuse, and analytics of suspicious behavior like native code and reflective call invocation. *AspectDroid* leverages the AspectJ instrumentation framework to inject monitoring code. The instrumented app is then executed dynamically to trace and log runtime activities at specific joinpoints. It can also instrument runtime classes for further analysis thus increasing code coverage. We have demonstrated that *AspectDroid* can achieve up to 94.68% F-score accuracy in detecting data leaks. Further analysis of 100 malware families for the Drebin dataset and 100 apps from Google Play showed our system can effectively analyze a diverse set of apps, including stealthy malware, with very minimal CPU and memory overhead.

Acknowledgment

This work was funded by the NSF grant, CNS #1409534.

REFERENCES

Afonso VM, de Amorim MF, Grégio ARA, Junquera GB, de Geus PL. Identifying Android malware using dynamically obtained features. J Comput Virol Hacking Tech 2015;11(1):9–17.

Afonso VM, de Geus PL, Bianchi A, Fratantonio Y, Kruegel C, Vigna G, et al. Going native: using a large-scale analysis of Android apps to create a practical native-code sandboxing policy. In NDSS; 2016.

Ali-Gombe A, Ahmed I, Richard GG III, Roussev V. AspectDroid: Android app analysis system. In Proceedings of the Sixth ACM on Conference on Data and Application Security and Privacy, pages 145–147. ACM, 2016.

Android Studio. Android developers; 2015. [Accessed 11 August 2015].

Apache Software Foundation. Apache HTTP Server Project. 2013. https://httpd.apache.org/download.cgi. [Accessed 11 December 2015].

Apache Software Foundation. Apache commons – common lang; 2015. https://commons.apache.org/proper/commons-lang/download_lang.cgi. [Accessed 30 August 2015].

APIMonitor. Installation and usage of DroidBox APIMonitor; 2012. [Accessed 6 May 2015].

Arp D, Spreitzenbarth M, Hübner M, Gascon H, Rieck K, CERT Siemens. Drebin: Effective and explainable detection of Android malware in your pocket. In Proceedings of the Annual Symposium on Network and Distributed System Security (NDSS), 2014.

AspectJ Team. The AspectJ TM programming guide; 2002–2003.

AspectJ Team. Eclipse – AspectJ compiler; 2015. [Accessed 26 August 2015].

Backes M, Gerling S, Hammer C, Maffei M, von Styp-Rekowsky P. Appguard–enforcing user requirements on Android apps. In: Tools and algorithms for the construction and analysis of systems. Springer; 2013. p. 543–8.

Bartel A, Klein J, Monperrus M, Allix K, Le Traon Y. Improving privacy on Android smartphones through in-vivo bytecode instrumentation. Technical Report 978-2-87971-111-9, uni. lu; 2012.

Blasing T, Batyuk L, Schmidt A-D, Camtepe SA, Albayrak S. An Android application sandbox system for suspicious software detection. In Malicious and unwanted software (MALWARE), 2010 5th international conference on, pages 55–62. IEEE; 2010.

Bob P. Dex2jar; 2014. [Accessed 13 December 2014].

Bruening Q, Zhao D, Amarasinghe S. Transparent dynamic instrumentation. In International Conference on Virtual Execution Environments, VEE-12; 2012.

Buck B, Hollingsworth JK. An API for runtime code patching. Int J High Perform Comput Appl 2000;14(4):317–29.
Burguera I, Zurutuza U, Nadjm-Tehrani S. Crowdroid: Behavior-based malware detection system for Android. In Proceedings of the 1st ACM Workshop on Security and Privacy in Smartphones and Mobile Devices, SPSM '11, pages 15–26; 2011.
DroidBox. Droidbox – Android application sandbox; 2011. [Accessed 6 May 2015].
DROIDBENCH. Secure software engineering at the European center for security and privacy by design – ec-spride; 2015. [Accessed 8 May 2015].
Enck W, Ongtang M, McDaniel P. On lightweight mobile phone application certification. In Proceedings of the 16th ACM Conference on Computer and Communications Security, CCS '09, pages 235–245; 2009.
Enck W, Gilbert P, Chun B-G, Cox LP, Jung J, McDaniel P, et al. TaintDroid: An information-flow tracking system for realtime privacy monitoring on smartphones. In Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation, OSDI'10; 2010.
Falcone Y, Currea S, Jaber M. Runtime verification and enforcement for android applications with RV-Droid. In: Runtime Verification, vol. 7687. Lecture Notes in Computer Science. 2013. p. 88–95.
Felt AP, Chin E, Hanna S, Song D, Wagner D. Android permissions demystified. In Proceedings of the 18th ACM Conference on Computer and Communications Security, CCS '11, pages 627–638; 2011.
Feng Y, Anand S, Dillig I, Aiken A. Apposcopy: Semantics-based detection of Android malware through static analysis. In Proceedings of the 22Nd ACM SIGSOFT International Symposium on Foundations of Software Engineering, FSE 2014, pages 576–587; 2014.
GDATA Software. GDATA mobile malware report: Q4/2015; 2016. https://public.gdatasoftware.com/Presse/Publikationen/Malware_Reports/US/G_DATA_MobileMWR_Q4_2015_US.pdf. [Accessed 3 October 2016].
Gibler C, Crussell J, Erickson J, Chen H. Androidleaks: Automatically detecting potential privacy leaks in android applications on a large scale. In: Trust and Trustworthy Computing, vol. 7344. Lecture Notes in Computer Science. 2012. p. 291–307.
Jeon J, Micinski KK, Vaughan JA, Fogel A, Reddy N, Foster JS, et al. Dr. Android and Mr. Hide: Fine-grained permissions in Android applications. Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices, SPSM '12, pages 3–14, New York, NY, USA; 2012. ACM.
Karami M, Elsabagh M, Najafiborazjani P, Stavrou A. Behavioral analysis of Android applications using automated instrumentation. In Software Security and Reliability-Companion (SERE-C), 2013 IEEE 7th International Conference on, pages 182–187. IEEE, 2013.
Kiczales G, Lamping J, Mendhekar A, Maeda C, Lopes C, Loingtier J-M, et al. Aspect-oriented programming. In: ECOOP'97 Object-oriented Programming, vol. 1241. Lecture Notes in Computer Science. 1997. p. 220–42.
Kiczales GJ, Lamping JO, Lopes CV, Hugunin JJ, Hilsdale EA, Boyapati C. Aspect-oriented programming, October 15 2002. US Patent 6,467,086; 2002.
Luk C-K, Cohn R, Muth R, Patil H, Klauser A, Lowney G, et al. Pin: building customized program analysis tools with dynamic instrumentation. In Proceedings of the 2005 ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '05, pages 190–200; 2005.
Mutti S, Fratantonio Y, Bianchi A, Invernizzi L, Corbetta J, Kirat D, et al. BareDroid: Large-scale analysis of Android apps on real devices. In Proceedings of the 31st Annual Computer Security Applications Conference, pages 71–80. ACM; 2015.
Rastogi V, Chen Y, Enck W. AppsPlayground: automatic security analysis of smartphone applications. In Proceedings of the Third ACM Conference on Data and Application Security and Privacy, CODASPY '13, pages 209–220; 2013.
Tam K, Khan SJ, Fattori A, Cavallaro L. CopperDroid: automatic reconstruction of Android malware behaviors. In NDSS; 2015.
Trend Micro. Android malware: how worried should you be?; 2012. [Accessed 4 October 2016].
Vasudevan A, Yerraballi R. SPiKE: engineering malware analysis tools using unobtrusive binary-instrumentation. In Proceedings of the 29th Australasian Computer Science Conference – Volume 48, ACSC '06, pages 311–320; 2006.
Weichselbaum L, Neugschwandtner M, Lindorfer M, Fratantonio Y, van der Veen V, Platzer C. Andrubis: Android malware under the magnifying glass. Vienna University of Technology, Tech. Rep. TRISECLAB-0414-001; 2014.
Wu D-J, Mao C-H, Wei T-E, Lee H-M, Wu K-P. DroidMat: Android malware detection through manifest and API calls tracing. In Proceedings of the 2012 Seventh Asia Joint Conference on Information Security, ASIAJCIS '12, pages 62–69; 2012.
Yan L-K, Yin H. DroidScope: Seamlessly reconstructing the OS and Dalvik semantic views for dynamic Android malware analysis. In USENIX Security Symposium, pages 569–584; 2012.
Zhang M, Yin H. Efficient, context-aware privacy leakage confinement for Android applications without firmware modding. In Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security, ASIA CCS '14, pages 259–270; 2014a.
Zhang M, Yin H. Appsealer: automatic generation of vulnerability-specific patches for preventing component hijacking attacks in Android applications. In Proceedings of the 21th Annual Network and Distributed System Security Symposium, NDSS 2014; 2014b.
Zhou Y, Jiang X. Dissecting Android malware: characterization and evolution. Security and Privacy (SP), 2012 IEEE Symposium on, pages 95–109; 2012.
Zhou Y, Wang Z, Zhou W, Jiang X. Hey, you, get off of my market: detecting malicious apps in official and alternative Android markets. In Proceedings of the Network and Distributed System Security Symposium, NDSS2012; 2012.

Aisha Ali-Gombe is an assistant professor of Computer Science at Towson University and a research scientist with the Center for Computation and Technology Louisiana State University. Dr. Ali-Gombe earned her Ph.D. in Engineering and Applied Science with major in Computer Science from the University of New Orleans in May 2017, following an M.S. degree in Computer Science in 2012. Her research interest in cybersecurity and digital forensics include code fingerprinting, malware analysis, privacy policy enforcement techniques, mobile security and memory and database forensics.

Brendan Saltaformaggio is an assistant professor in the School of Electrical and Computer Engineering at Georgia Tech, with a courtesy appointment in the School of Computer Science. His research interests lie in computer systems security, cyber forensics, and the vetting of untrusted software. Dr. Saltaformaggio serves as the Director of the Cyber Forensics Innovation (CyFI) Laboratory. The CyFI Lab's mission is to further the investigation of advanced cyber crimes and the analysis and prevention of next-generation malware attacks, particularly in mobile and IoT environments. This research has led to numerous publications at top cyber security venues, including a Best Paper Award from the ACM Conference on Computer and Communications Security (CCS'15) and a Best Student Paper Award from the 2014 USENIX Security Symposium.

Originally from New Orleans, Dr. Saltaformaggio earned his Bachelor of Science with Honors in Computer Science from the University of New Orleans in 2012. He received his M.S. and Ph.D. in Computer Science at Purdue University in 2014 and 2016, respectively, during which Dr. Saltaformaggio was honored with the 2017 ACM SIGSAC Doctoral Dissertation Award as well as two fellowships: the 2016 Symantec Research Labs Graduate Fellowship and the inaugural Emil Stefanov Memorial Fellowship in Computer Science.

J. "Ram" Ramanujam is a Floating-Point Systems Endowed Chair in Computational Methods; John E. & Beatrice L. Ritter distinguished professor of Electrical and Computer Engineering and the Director of Center for Computation and Technology, Louisiana State University. Dr. Ramanujan earned his Ph.D. from the Ohio State in 1990, and since then has been a faculty member at LSU. His research interests are in compilers and runtime systems for high-performance computing, domain-specific languages, compilers for parallel computing, embedded systems and energy-aware computing systems and cybersecurity in mobile and control systems.

Dongyan Xu is a professor of Computer Science at Purdue University. He is also the interim director of the Center for Education and Research in Information Assurance and Security (CERIAS). He has been on Purdue faculty since 2001, when he received his Ph.D. in Computer Science from the University of Illinois at Urbana-Champaign. His research efforts span computer systems security and forensics, cloud computing, and virtualization, with projects sponsored by both government agencies and industry. He is the co-author of six award-winning papers at major conferences in security and cloud computing.

Golden G. Richard III is a professor of Computer Science and Engineering and Associate Director for Cybersecurity at the Center for Computation and Technology at Louisiana State University. Dr. Richard is also a Fellow of the American Academy of Forensic Sciences (AAFS). He earned a B.S. in Computer Science (with honors) from the University of New Orleans and an M.S. and Ph.D. from the Ohio State University. His research interests include digital forensics, reverse engineering, offensive computing, operating systems internals, and malware analysis.